

UNITED STATES DEPARTMENT OF COMMERCE United States Patent and Trademark Office Address COMMISSIONER FOR PATENTS PO Box 1450 Alexascins, Virginia 22313-1450 www.emplo.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/693,227	10/23/2003	Jerome R. Bellegarda	4860P3128	2262
8791 7590 080772008 BLAKELY SOKOLOFF TAYLOR & ZAFMAN LLP 1279 OAKMEAD PARKWAY			EXAMINER	
			RIDER, JUSTIN W	
SUNNYVALE, CA 94085-4040		ART UNIT	PAPER NUMBER	
			2626	
			MAIL DATE	DELIVERY MODE
			08/07/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Application No. Applicant(s) 10/693 227 BELLEGARDA, JEROME R. Office Action Summary Examiner Art Unit JUSTIN W. RIDER 2626 -- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --Period for Reply A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS. WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION. Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication. If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication - Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b). Status 1) Responsive to communication(s) filed on 25 April 2008. 2a) ☐ This action is FINAL. 2b) This action is non-final. 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under Ex parte Quayle, 1935 C.D. 11, 453 O.G. 213. Disposition of Claims 4) Claim(s) See Continuation Sheet is/are pending in the application. 4a) Of the above claim(s) is/are withdrawn from consideration. 5) Claim(s) _____ is/are allowed. 6) ☐ Claim(s) See Continuation Sheet is/are rejected. 7) Claim(s) 16, 40, 64 and 88 is/are objected to. 8) Claim(s) _____ are subject to restriction and/or election requirement. Application Papers 9) The specification is objected to by the Examiner. 10) The drawing(s) filed on is/are; a) accepted or b) objected to by the Examiner. Applicant may not request that any objection to the drawing(s) be held in abevance. See 37 CFR 1.85(a). Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d). 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152. Priority under 35 U.S.C. § 119 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f). a) All b) Some * c) None of: Certified copies of the priority documents have been received. 2. Certified copies of the priority documents have been received in Application No. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)). * See the attached detailed Office action for a list of the certified copies not received. Attachment(s) 1) Notice of References Cited (PTO-892) 4) Interview Summary (PTO-413) Paper No(s)/Mail Date. ___ Notice of Draftsperson's Patent Drawing Review (PTO-948) 5) Notice of Informal Patent Application 3) Information Disclosure Statement(s) (PTO/SB/08)

Paper No(s)/Mail Date _

6) Other:

Application No. 10/693,227

Continuation of Disposition of Claims: Claims pending in the application are 1, 4-25, 27-49, 51-73, 75-97, 99, 101, 35 U.S.C. § 103, 105, 107, 109 and 111.

Application/Control Number: 10/693,227 Page 2

Art Unit: 2626

Continued Examination Under 37 CFR 1.114

 A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 25 April 2008 has been entered.

Response to Arguments

2. Applicant's arguments filed 25 April 2008 have been fully considered but they are not persuasive. On pages 28-29 of Remarks, with respect to the teachings of COORMAN, applicant appears to merely be pointing out the motivation for using BANBROOK to clearly show obviousness of claimed limitations and so therefore is moot.

Further, on page 29 of Remarks, applicant appears to traverse the teachings of BANBROOK regarding the construction (e.g., creation, production, etc) and subsequent decomposition of a matrix. As stated within applicant's remarks, 'BANBROOK discloses "the method of a singular value decomposition (SVD)..." as well as 'BANBROOK discloses producing the...matrix...' BANBROOK produces the matrices in response to gathered speech data. This, in combination with COORMAN teaches the limitations as claimed.

Applicant further limited the claims by more clearly defining portions as 'portions surrounding the segment boundary within a phoneme.' As taught by COORMAN, in columns 18 and 19, diphones are utilized as segments, or portions. In the Glossary section of COORMAN (cols. 19-20), diphones are defined as a fundamental speech unit composed of two

Art Unit: 2626

adjacent half-phones. Thus, the left and right boundaries of a diphone are in-between phone boundaries. The center of the diphone contains the phone-transition region. From this definition, it appears that COORMAN clearly extracts portions having boundaries within a given phoneme.

Claim Rejections - 35 USC § 103

- The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:
 - (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.
- (It is noted that the limitations below are grouped together for sake of length and that they are identical but for the statutory categories they represent.)
- 4. Claims 1, 3-8, 19-20, 25, 27-32, 43-44, 49, 51-56, 67-68, 73, 75-80, 91-92, 97, 99, 101, 103, 105, 107, 109 and 111 are rejected under 35 U.S.C. 103(a) as being unpatentable over COORMAN in view of Michael BANBROOK, 'Nonlinear Analysis of Speech From a Synthesis Perspective', A thesis submitted for the degree of Doctor of Philosophy at The University of Edinburgh; October 15, 1996 (Specifically Chapter 4), referred to as BANBROOK hereinafter.
- Claims 1, 25, 49 and 73: COORMAN discloses a method (as disclosed in COORMAN see claims 51-93), computer-readable medium having instructions to cause a machine to perform a machine-implemented method (col. 19, lines 20-30), apparatus (i.e., speech synthesizer, see

Art Unit: 2626

Claims 1-50) and a system (see col. 19, 'Conclusion') comprising a processing unit (e.g., 'Text Processor') coupled to a memory through a bus (see col. 19, 'Conclusion', which describes the flexibility of the system to be implemented on any number of computer systems using a plurality of either <u>fixed</u> of removable storage means (e.g., memory).) for analyzing speech for use in synthesis, comprising:

- i. extracting portions from time-domain speech segments (col. 5, lines 28-30), the portions surrounding a segment boundary within a phoneme (col. 19, lines 1-9);
- ii. creating feature vectors (col. 5, lines 28-30) that represent the portions in a vector space; and
- determining a distance between the feature vectors in the vector space (col. 18, lines 16-19).

However, COORMAN fails to, but BANBROOK does specifically disclose wherein features include phase information of the portions (p. 37, 'The data is projected onto a phase space defined by the singular vectors of the data, which can then be partitioned into a signal subspace and a noise subspace.') and wherein creating feature vectors comprises constructing a matrix W from the portions (col. 18, lines 21-23, 'The calculation of this spectral mismatch is based on a distance calculation between spectral vectors. This might be a heavy task as there can be many segment combinations possible. In order to reduce the computational complexity a combination matrix—containing the spectral distances—could be calculated in advance.' [emphasis added]).

However, COORMAN recites performing operations on said matrix containing the portions surrounding the segment boundary within the phoneme (i.e., diphone), however failing

Art Unit: 2626

to, but BANBROOK does specifically disclose decomposing the matrix W (p. 37, 'The method of singular value decomposition (SVD) reduction, described by Broomhead and King [85, 103], addresses this problem.').

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of BANBROOK in the system of COORMAN because it introduces a combination of analysis tools (e.g. time delay embedding, singular value decomposition, correlation dimension, local singular value analysis, Lyapunov spectra and short term prediction properties) and looks in detail at Lyapunov exponents and two major novel modifications are proposed that are demonstrated to be more robust than conventional techniques (Abstract).

Claims 3, 27, 51 and 75: COORMAN discloses a system as per claims 1, 25, 49 and 73 above, further comprising extracting global boundary-centric features from the portions (col. 10, lines 49-54).

Claims 4, 28, 52 and 76: COORMAN discloses a system as per claims 1, 25, 49 and 73 above, wherein the speech segments each include the segment boundary within the phoneme (col. 9, lines 5-8).

<u>Claims 5, 29, 53 and 77</u>: COORMAN discloses a system as per claims 4, 28, 52 and 76 above, wherein the speech segments each include at least one diphone (col. 9, lines 5-8).

Claims 6, 30, 54 and 78: COORMAN discloses a system as per claims 5, 29, 53 and 77 above, wherein the portions include at least one pitch period (col. 19, lines 7-9).

Claims 7, 31, 55 and 79: COORMAN, in view of BANBROOK disclose a system as per claims 6, 30, 54 and 78 above. However, COORMAN fails to, but BANBROOK does

Art Unit: 2626

specifically disclose wherein decomposing the matrix W comprises performing a pitch synchronous (p. 37, 'which can then be partitioned into a signal subspace and a noise subspace.') singular value analysis on the pitch periods of the time-domain segments (p. 37, 'The method of singular value decomposition (SVD) reduction, described by Broomhead and King [85, 103], addresses this problem.').

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of BANBROOK in the system of COORMAN because of the reasons described above.

Claims 8, 32, 56 and 80: COORMAN, in view of BANBROOK disclose a system as per claims 6, 30, 54 and 78 above. However, COORMAN fails to, but BANBROOK does specifically disclose wherein the matrix W is a 2KMxN matrix represented by

$$W = U\Sigma V^T$$

where K is the number of pitch periods near the segment boundary extracted from each segment, N is the maximum number of samples among the pitch periods, M is the number of segments in a voice table having a segment boundary within the phoneme, U is the $2KM \times R$ (p. 37, N x w trajectory matrix found utilizing time delay embedding) left singular matrix with row vectors u_i ($l \le i \le 2KM$), Σ is the $R \times R$ diagonal matrix of singular values $s_1 \ge s_2 \ge ... \ge s_R > 0$, V is the $N \times R$ right singular matrix with row vectors v_i ($l \le j \le N$), R << 2KM, and T denotes matrix transposition, wherein decomposing the matrix W comprises performing a singular value decomposition of W (p. 37-38, $X = S\Sigma C^T$, where X is the trajectory matrix, S and C are the matrices of the singular vectors associated with Σ , which is a diagonal matrix.).

Art Unit: 2626

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of BANBROOK in the system of COORMAN because of the reasons described above.

Claims 19, 43, 67 and 91: COORMAN discloses a system as per claims 5, 29, 53 and 77 above, wherein the portions include centered pitch periods (col. 19, lines 7-9, 'In the preferred embodiment the length of the trailing and leading regions are of the order of one to two pitch periods and the sliding window is bell-shaped fi.e. centered].').

Claims 17-18, 41-42, 65-66 and 89-90: COORMAN discloses a system as per claims 1, 25, 49 and 73 above, wherein said distances are associated with said speech segments (units, col. 11, section 'Cost Functions', lines 46-49, 'a set of nonlinear cost functions has been defined for use in the unit selection...with specific properties which help in the unit selection process.').

Claims 20, 44, 68 and 92: Claims 20, 44, 68 and 92 are similar in scope and content to that of claim 8 above and so therefore are rejected under the same rationale.

Claims 97, 101, 105 and 109: COORMAN discloses a method (as disclosed in COORMAN see claims 51-93), computer-readable medium having instructions to cause a machine to perform a machine-implemented method (col. 19, lines 20-30), an apparatus (i.e., speech synthesizer, see Claims 1-50) and a system (see col. 19, 'Conclusion') comprising a processing unit (e.g., 'Text Processor') coupled to a memory through a bus (see col. 19, 'Conclusion', which describes the flexibility of the system to be implemented on any number of computer systems using a plurality of either fixed of removable storage means (e.g., memory).) for analyzing speech for use in synthesis, comprising:

Art Unit: 2626

i. gathering time-domain samples from recorded speech segments (col. 20, lines 17-22,

'The database may directly contain <u>digitally sampled waveforms</u>, or it may include pointers to

such waveforms,' [emphasis supplied]) wherein the time-domain samples include pitch periods
surrounding a segment boundary within a phoneme (col. 19, lines 1-9);

ii. extracting features that represent the samples (col. 4, lines 23-25, 'The acoustic join cost is based on a quantization of the mel-cepstrum,');

iii. determining a discontinuity between the segments (col. 12, 'Cost Functions for Numeric Features', 'Imprecise linguistic or acoustic knowledge, for example, how big a discontinuity in pitch can be perceived,'), the discontinuity based on a distance between the features ('For example, the mismatch of pitch between phones with the same accentuation (either both accented, or both unaccented) in the Transition Cost has a symmetric cost function...').

However, COORMAN fails to, but BANBROOK does specifically disclose wherein features include phase information of the portions (p. 37, 'The data is projected onto a phase space defined by the singular vectors of the data, which can then be partitioned into a signal subspace and a noise subspace.') and wherein creating feature vectors comprises constructing a matrix W from the portions (col. 18, lines 21-23, 'The calculation of this spectral mismatch is based on a distance calculation between spectral vectors. This might be a heavy task as there can be many segment combinations possible. In order to reduce the computational complexity a combination matrix—containing the spectral distances—could be calculated in advance.' [emphasis added]).

However, COORMAN recites performing operations on said matrix containing the time domain samples of the pitch periods surrounding the segment boundary, however failing to, but

Art Unit: 2626

BANBROOK does specifically disclose decomposing the matrix W (p. 37, 'The method of singular value decomposition (SVD) reduction, described by Broomhead and King [85, 103], addresses this problem.').

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of BANBROOK in the system of COORMAN because it introduces a combination of analysis tools (e.g. time delay embedding, singular value decomposition, correlation dimension, local singular value analysis, Lyapunov spectra and short term prediction properties) and looks in detail at Lyapunov exponents and two major novel modifications are proposed that are demonstrated to be more robust than conventional techniques (Abstract).

Claims 99, 103, 107 and 111: COORMAN discloses a system as per claims 98, 102, 106, and 110 above. However, COORMAN fails to, but BANBROOK does specifically disclose wherein features include phase information of the portions (p. 37, 'The data is projected onto a phase space defined by the singular vectors of the data, which can then be partitioned into a signal subspace and a noise subspace.').

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of BANBROOK in the system of COORMAN because of the reasons described above.

Claims 9-10, 21-23, 33-34, 45-47, 57-58, 69-71, 81-82 and 93-95 are rejected under 35
 U.S.C. 103(a) as being unpatentable over COORMAN, in view of BANBROOK and in further

Art Unit: 2626

view of ANSARI et al., 'Pitch Modification of Speech Using a Low-Sensitivity Inverse Filter

Approach'; IEEE Signal Processing Letters; March 1998 referred to as ANSARI hereinafter.

Claims 9, 33, 57 and 81: COORMAN, in view of BANBROOK disclose a system as per claims 8, 32, 56 and 80 above, however failing to, but ANSARI does specifically disclose padding a signal with zeroes (p. 61, section III, 'in the new method when the residual is modified with zero-padding to lower the pitch.').

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of ANSARI in the system of COORMAN, in view of BANBROOK because speech modifications using the method of ANSARI are superior in quality to those obtained with RELP, while at the same time being less sensitive than RELP to errors in pitch marking (Abstract).

Claims 10, 34, 58 and 82: COORMAN, in view of BANBROOK disclose a system as per claims 9, 33, 57 and 81 above. However, COORMAN fails to, but BANBROOK does specifically disclose wherein a feature vector u_i is calculated as

$$\bar{u}_i = u_i \Sigma$$

where u_l is a row vector associated with a pitch period i, and Σ is the singular diagonal matrix (p.49, 'In general, any matrix A can be written A = QR (4.17) where Q has orthogonal columns and R is a square upper-right triangular matrix with positive values on the diagonal.').

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of BANBROOK in the system of COORMAN because of the reasons described above.

Art Unit: 2626

Claims 21, 45, 69 and 93: COORMAN, in view of BANBROOK disclose a system as per claims 20, 44, 68 and 92 above, however failing to, but ANSARI does specifically disclose symmetrically padding a signal with zeroes (p. 61, section III, 'in the new method when the residual is modified with zero-padding to lower the pitch.'). It would have been obvious to one having ordinary skill in the art that if pitch periods were centered, that one would be motivated to append zeros symmetrically on either side of the centered samples in order to maintain symmetric proportions with respect to a centered pitch.

Claims 22-23, 46-47, 70-71 and 94-95: Claims 22-23, 46-47, 70-71 and 94-95 are similar in scope and content to that of claims 10 and 12 above and so therefore are rejected under the same rationale.

6. Claims 11-15, 35-39, 59-63 and 83-86 are rejected under 35 U.S.C. 103(a) as being unpatentable over COORMAN and BANBROOK in view of ANSARI and in further view of Jerome R. BELLEGARDA, 'Exploiting Latent Information in Statistical Language Modeling' referred to as BELLEGARDA hereinafter.

Claims 11, 35, 59 and 83: COORMAN and BANBROOK in view of ANSARI disclose a system as per claims 10, 34, 58 and 82 above. However, COORMAN and BANBROOK in view of ANSARI fail to, but BELLEGARDA does specifically disclose wherein the distance between two feature vectors is determined by a metric comprising the cosine of the angle between the two feature vectors (p. 5, 'We conclude that a natural metric to consider for the "closeness" between words is therefore the cosine of the angle between ū₁ and ū₁.).

Art Unit: 2626

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of BELLEGARDA in the system of COORMAN and BANBROOK in view of ANSARI because it uses latent semantic analysis to improve statistical language modeling utilizing existing clustering techniques capable of aiding in speech production by machine (Introduction).

Claims 12, 36, 60 and 84: COORMAN and BANBROOK in view of ANSARI disclose a system as per claims 10, 34, 58 and 82 above. However, COORMAN and BANBROOK in view of ANSARI fail to, but BELLEGARDA does specifically disclose wherein the metric comprises a closeness measure, C, between two feature vectors, \bar{u}_k and \bar{u}_l , wherein C is

$$\zeta(\vec{u}_k, \vec{u}_l) = \cos(u_k \Sigma, u_l \Sigma) = \frac{u_k \Sigma^2 u_l^T}{\|u_k \Sigma\| \|u_l \Sigma\|}$$
 calculated as

for any $1 \le k$, $l \le 2KM$ (p. 6, (10)).

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of BELLEGARDA in the system of COORMAN and BANBROOK in view of ANSARI because of the reasons described above.

Claims 13, 37, 61 and 85: COORMAN and BANBROOK in view of ANSARI and in further view of BELLEGARDA disclose a system as per claims 12, 36, 60 and 84 above. The

$$d(S_1,S_2) = d_0(p_1,q_1) = 1 - C(\widetilde{u}_{\rho_1},\widetilde{u}_{\varrho_2})$$

examiner is taking Official Notice that the difference as calculated by

is simply a natural extension from the closeness measure as determined in the prior claim (which is assumed a value between 0 and 1). Therefore, it would have been obvious to one having

Art Unit: 2626

ordinary skill in the art at the time of invention to include a difference measure as the counterpart to a closeness measure previously determined because it is well known that the two factors have an inversely variable relationship, here adding up to 1.

Claims 14, 38, 62 and 86: COORMAN discloses a system as per claims 13, 37, 61 and 85 above, wherein the calculation for the difference between two segments in the voice table, S1 and S2, is expanded to include a plurality of pitch periods from each segment (col. 19, lines 1-9).

Claims 15, 39, 63 and 87: COORMAN discloses a system as per claims 13, 37, 61 and 85 above, wherein the difference between two segments in the voice table, S₁ and S₂, is associated with a discontinuity between S₁ and S₂ (col. 18, lines 48-54, 'The major concern of waveform concatenation is in avoiding waveform irregularities such as discontinuities and fast transients that may occur in the neighborhood of the join...It is thus important to minimize signal discontinuities at each junction.').

Allowable Subject Matter

7. Claims 16, 40, 64 and 88 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

The above claims recite a specific distance formula as follows:

$$d(S_1,S_2) = \left| d_0(p_1,q_1) - \underline{d_0(p_1,\overline{p}_1) + d_0(q_1,\overline{q}_1)} \right| = \left| \underline{C(\overline{u}_{\theta^1},\overline{u}_{\theta^1}) + C(\overline{u}_{\theta^1},\overline{u}_{\theta^1})} - C(\overline{u}_{\theta^1},\overline{u}_{\theta^1}) \right|$$

The above is used as an alternative distance measure that is essentially the relative change in similarity that occurs during a concatenation function. More specifically, this alternative distance specifically shows wherein a difference is zero only when two identical segments are

Art Unit: 2626

concatenated together; otherwise a difference measure greater than zero exists. While the cited prior art references do use distance measures as disclosed, none of the references use an alternative distance measure as specifically disclosed and defined as per claims 16, 40, 64 and 88.

Conclusion

 Any inquiry concerning this communication or earlier communications from the examiner should be directed to JUSTIN W. RIDER whose telephone number is (571)270-1068.
 The examiner can normally be reached on Monday - Friday 8:30AM - 5:00PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/J. W. R./ Examiner, Art Unit 2626 31 July 2008 /David R Hudspeth/ Supervisory Patent Examiner, Art Unit 2626